Causal Effect of Caffeine on Heart Rate

using latin square design

Young Geun Kim
ygeunkim.github.io
2019711358, Department of Statistics

20 May, 2021

Abstract

After drinking coffee, one's heartbeat raises. This kind of change in metabolism might be responsible for caffeine. To confirm this, we conduct an experiment. Blocked by coffe-to-water ratio and drinking speed, we build 4×4 latin square design with caffeine intensity factor. For each cell, we measure heart rate using ECG app installed in Apple Watch (it gives average BPM).

Contents

1 Introduction

We try to experiment whether caffeine affect ECG or heart rate

- Caffeine through coffee
- Average heart rate (BPM)

2 Design of the Experiment

2.1 Latin Square Design

In this experiment, we implement reduced latin square design (LSD).

2.1.1 Blocking factors

It is known that caffeine intake depends on coffee-to-water ratio and drinking speed (한동하 2018). So we set these two as row and column factors.

Coffee-to-water ratio is our row factor.

1. 1:0 (Espresso, 40 ml)
2. 1:2.5 (Water 100 ml)
3. 1:5 (Water 200 ml)
4. 1:7.5 (Water 300 ml)

Drinking-coffee-speed is our column speed.

1. <=5 sec
2. 5-15 sec
3. 15-30 sec
4. 30< sec

2.1.2 Treatment

Denote that we are interested in caffeine. We took capsule, Starbucks by Nespresso (himynameisabcde 2020).

1. House blend: 74.5 mg
2. Sumatra: 54.5 mg
3. Decaf espresso roast: 3 mg
4. Just water: 0 mg

2.1.3 Output and measurement

Using ECG app in Apple Watch series 4, we measure *Average heart rage (in BPM)*. Since there exists variation, we consider difference of BPM after and before: $y_{rc}^{post} - y_{rc}^{pre}$

2.1.4 Remarks

First we randomly allocate treatments (`1`, `2`, `3`, `4`) to (`A`, `B`, `C`, `D`).

```
set.seed(1)
sample(LETTERS[1:4])
#> [1] "A" "C" "D" "B"
```

i.e.

(A) House blend (74.5 mg)

(B) Water (0 mg)

(C) Sumatra (54.5 mg)

(D) Decaf espresso roast (3 mg)

Table 1 shows the final design.

Table 1: Design of the Experiment

	Drinking Speed			
Water	<=5	5-15	15-30	30<
0 ml	H(74.5)	W(0)	S(54.5)	D(3)
100 ml	W(0)	S(54.5)	D(3)	H(74.5)
200 ml	S(54.5)	D(3)	H(74.5)	W(0)
300 ml	D(3)	H(74.5)	W(0)	H(74.5)

[1] 'Water' is the coffee-to-water ratio (divide with 40 ml)

[2] Numbers in the brackets indicate caffeine (in mg)

We try to control the other variables as possible as we can. When taking coffee,

- Drink coffee every morning (between 8:30 a.m. and 9:00 a.m.)
- after eating a piece of bread
- Nespresso machine: Pixie C61 in my home

When measuring heart rate,

- Sitting at the table
- Rest my arms on the table
- Use the same strip
  - Nike sport band
  - of same fit (8-th)
- and other instructions in https://support.apple.com/en-us/HT208955

3 Data Analysis

3.1 Data

See Table 2 for the result of the experiment.

Table 2: Experiment Data

	Drinking Speed			
water	<=5	5-15	15-30	30<
0 ml	HB, 9	W, 1	S, 7	D, 3
100 ml	W, 2	S, 6	D, 3	HB, 14
200 ml	S, 4	D, 3	HB, 1	W, 0
300 ml	D, 2	HB, 4	W, 2	S, 4

[1] Caffeine: HB > S > D > W

[2] Numbers indicate the difference after and before taking coffee

3.2 ANOVA

Now we build ANOVA table. See 3.

Table 3: ANOVA Table

Source	Observed			F-Statistic
	DF	Sum Sq	Mean Sq	
water	3	44.19	14.73	1.625
speed	3	9.69	3.23	0.356
coffee	3	78.69	26.23	2.894
Residuals	6	54.38	9.06	
Total	15	186.94		

Recall that only $F_{Tre} = 2.894$ has causal meaning. Now we randomize F_{Tre} under sharp null.

3.3 Randomization Test

3.3.1 Sharp Null of No Effect

Consider the sharp null of no effect.

$$H_0 : y_{rc}(1) = y_{rc}(2) = y_{rc}(3) = y_{rc}(4) \tag{1}$$

See Science Table in 4.

Table 4: Observed Values of the Science Table for the Coffee-ECG Experiment (row 2 and 3 omitted)

				Observed $y_{rc}(k)$			
id	water	speed	coffee	HB	W	S	De
Row 1 (Water 0 ml)							
1	1	1	1	9			
2	1	2	4		1		
3	1	3	2			7	
4	1	4	3				3
Row 2 (Water 100 ml)							
5	2	1	4		2		
6	2	2	2			6	
7	2	3	3				3
8	2	4	1	14			
Row 3 (Water 200 ml)							
9	3	1	2			4	
10	3	2	3				3
11	3	3	1	1			
12	3	4	4		0		
Row 4 (Water 300 ml)							
13	4	1	3				2
14	4	2	1	4			
15	4	3	4		2		
16	4	4	2			4	

Now we impute impute the missing $Y_{rc}(k)$ under the sharp null. See Table 5.

Table 5: Imputed Outcomes under the Sharp Null

				Observed $y_{rc}(k)$			
id	water	speed	coffee	HB	W	S	De

Row 1 (Water 0 ml)							
1	1	1	1	9	9	9	9
2	1	2	4	1	1	1	1
3	1	3	2	7	7	7	7
4	1	4	3	3	3	3	3
Row 2 (Water 100 ml)							
5	2	1	4	2	2	2	2
6	2	2	2	6	6	6	6
7	2	3	3	3	3	3	3
8	2	4	1	14	14	14	14
Row 3 (Water 200 ml)							
9	3	1	2	4	4	4	4
10	3	2	3	3	3	3	3
11	3	3	1	1	1	1	1
12	3	4	4	0	0	0	0
Row 4 (Water 300 ml)							
13	4	1	3	2	2	2	2
14	4	2	1	4	4	4	4
15	4	3	4	2	2	2	2
16	4	4	2	4	4	4	4

3.3.2 Randomization Test

Iterating 2000 times, p-value is

$$pvalue = 0.108$$

It is not significant as in ANOVA table (p-value of 0.124). Figure 1 presents the randomization distribution.

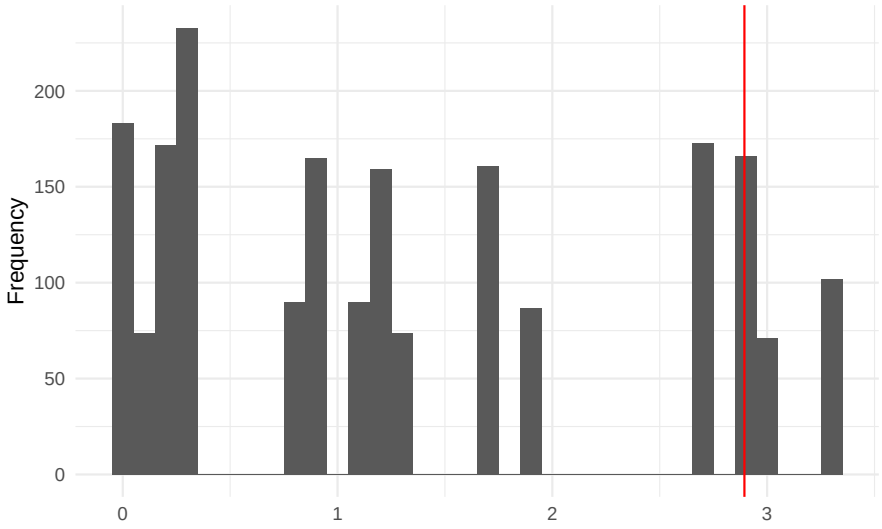


Figure 1: Randomization Distribution of F_{Tre} under the Null

4 Conclusion

4.1 Conclusion

We expected that caffeine effect was significant. However, there was *no significant evidence* (p-value of 0.108). Why?

- Caffeine **tolerance**
  - I have taken coffee everyday
  - Was coffee I have taken too small?
- **Outliers**
  - Unit `8` seems outliers
  - Table 2: value of 14

4.2 Future Study

- Values for the treatment were quite arbitrary, so we can re-try with other sets
- Re-measure (for Unit `8`)
  - Figure 2: Change the value of Unit `8` to 7
  - P-value becomes 0.046

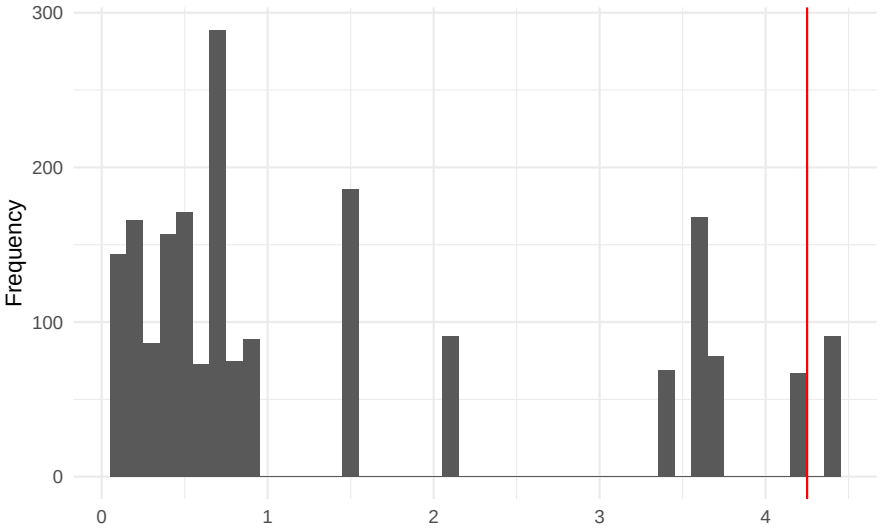


Figure 2: Changing the Outlier

References

himynameisabcde. 2020. “R/Nespresso - i Recieved the Caffiene Content Numbers for Starbucks Nespresso Pods!” *Reddit*. https://www.reddit.com/r/nespresso/comments/id31r5/i_recieved_the_caffiene_content_numbers_for/.

한동하. 2018. “[한동하 원장의 웰빙의 역설] 냉커피는 뜨거운 커피와 어떤 차이가 있을까?” *헬스경향*. http://www.k-health.com/news/articleView.html?idxno=37375.

A Appendix

A.1 Codes

Package loading:

```
# tidyverse family--------------------
library(tidyverse)
# kable-------------------------------
library(knitr)
library(kableExtra)
# LSD randomization-------------------
library(magic)
# permutation-------------------------
library(combinat)
# parallelization---------------------
library(foreach)
# set seed for report -----------------
set.seed(1)
# functions---------------------------
source("lsd-anova.R")
source("lsd-rand.R")
```

Design:

```
# tibble------------------------
latin <-
  tibble(
    row = rep(1:4, each = 4),
    col = rep(1:4, 4),
    trt = c(LETTERS[1:4], c("B", "C", "D", "A"), c("C", "D", "A", "B"), c("D", "A", "B", "A"))
  )
# kable-------------------------------
latin %>%
  transmute(
    Water = case_when(
      row == 1 ~ "0 ml",
      row == 2 ~ "100 ml",
      row == 3 ~ "200 ml",
      row == 4 ~ "300 ml"
    ),
    speed = case_when(
      col == 1 ~ "<=5",
      col == 2 ~ "5-15",
      col == 3 ~ "15-30",
      col == 4 ~ "30<"
    ),
    coffee = case_when(
      trt == "A" ~ "H(74.5)",
      trt == "B" ~ "W(0)",
      trt == "C" ~ "S(54.5)",
      trt == "D" ~ "D(3)"
    )
  ) %>%
  group_by(Water) %>%
  pivot_wider(names_from = speed, values_from = coffee) %>%
```

```
  kable(
    format = kable_format,
    escape = FALSE,
    caption = "Design of the Experiment"
  ) %>%
  kable_styling(full_width = FALSE, latex_options = c("striped", "HOLD_position")) %>%
  add_header_above(
    header = c(" " = 1, "Drinking Speed" = 4),
    escape = FALSE
  ) %>%
  footnote(
    number = c(
      "'Water' is the coffee-to-water ratio (divide with 40 ml)",
      "Numbers in the brackets indicate caffeine (in mg)"
    )
  )
```

Data:

```
# ANOVA-------------------------------------
ecg_anova <- aov_lsd(
  diff ~ water + speed + coffee,
  data = ecg_diff
)
ecg_aov <- broom::tidy(ecg_anova)
# Print ANOVA-------------------------------
print_anova(ecg_aov, format = kable_format, caption = "ANOVA Table") %>%
  row_spec(3, color = "red")
```

ANOVA:

```
# ANOVA-------------------------------------
ecg_anova <- aov_lsd(
  diff ~ water + speed + coffee,
  data = ecg_diff
)
ecg_aov <- broom::tidy(ecg_anova)
# Print ANOVA-------------------------------
print_anova(ecg_aov, format = kable_format, caption = "ANOVA Table") %>%
  row_spec(3, color = "red")
```

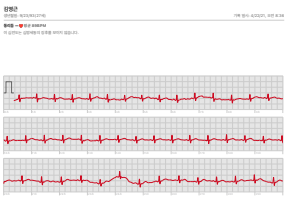
Randomization:

```
# science table-----------------------------------------------------------
ecg_po <- read_csv("../data/processed/science.csv")
ecg_po %>%
  # slice(1:4, 13:16) %>%
  kable(
    format = kable_format,
    escape = FALSE,
    caption = "Observed Values of the Science Table for the Coffee-ECG Experiment (row 2 and 3 omitted)
  ) %>%
  kable_styling(full_width = FALSE, latex_options = c("striped", "HOLD_position")) %>%
  add_header_above(
    header = c(" " = 4, "Observed $y_{rc}(k)$" = 4),
    escape = FALSE
```

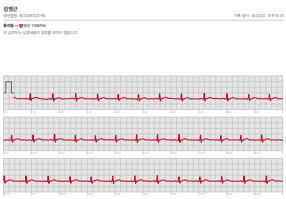
```
  ) %>%
  pack_rows(
    "Row 1 (Water 0 ml)",
    1, 4
  ) %>%
  pack_rows(
    "Row 2 (Water 100 ml)",
    5, 8
  ) %>%
  pack_rows(
    "Row 3 (Water 200 ml)",
    9, 12
  ) %>%
  pack_rows(
    "Row 4 (Water 300 ml)",
    13, 16
  )
# imputing-------------------------------------------------------------------
ecg_sharp <- read_csv("../data/processed/science-sharp.csv")
ecg_sharp %>%
  # slice(1:4, 13:16) %>%
  kable(
    format = kable_format,
    escape = FALSE,
    caption = "Imputed Outcomes under the Sharp Null",
    longtable = TRUE
  ) %>%
  kable_styling(full_width = FALSE, latex_options = c("striped", "HOLD_position")) %>%
  add_header_above(
    header = c(" " = 4, "Observed $y_{rc}(k)$" = 4),
    escape = FALSE
  ) %>%
  pack_rows(
    "Row 1 (Water 0 ml)",
    1, 4
  ) %>%
  pack_rows(
    "Row 2 (Water 100 ml)",
    5, 8
  ) %>%
  pack_rows(
    "Row 3 (Water 200 ml)",
    9, 12
  ) %>%
  pack_rows(
    "Row 4 (Water 300 ml)",
    13, 16
  )
# randomization test-------------------------------------------
num_iter <- 2000
f_tre <- test_sharp(ecg_sharp, ecg_diff$diff, num_iter, 4, c("water", "speed", "coffee"), 8)
p_val <- sum(f_tre >= ecg_aov$statistic[3]) / num_iter
# randomization distribution-----------------------------------
```

```r
tibble(f_stat = f_tre) %>%
  ggplot(aes(x = f_stat)) +
  geom_histogram(binwidth = .1) +
  geom_vline(xintercept = ecg_aov$statistic[3], col = I("red")) +
  theme_minimal() +
  labs(
    x = element_blank(),
    y = "Frequency"
  )
# Toy-----------------------------------
new_val <- 7
ecg_hb <-
  ecg_diff %>%
  mutate(diff = replace(diff, water == 2 & speed == 4 & coffee == 1, new_val))
sharp_hb <-
  ecg_sharp %>%
  mutate_at(vars(HB, W, S, De), ~replace(., water == 2 & speed == 4 & coffee == 1, new_val))
# anova with changed value----------------
hb_anova <- aov_lsd(
  diff ~ water + speed + coffee,
  data = ecg_hb
)
hb_aov <- broom::tidy(hb_anova)
# randomization with changed value--------
hb_tre <- test_sharp(sharp_hb, ecg_hb$diff, num_iter, 4, c("water", "speed", "coffee"), 8)
hb_pval <- sum(hb_tre >= hb_aov$statistic[3]) / num_iter
tibble(f_stat = hb_tre) %>%
  ggplot(aes(x = f_stat)) +
  geom_histogram(binwidth = .1) +
  geom_vline(xintercept = hb_aov$statistic[3], col = I("red")) +
  theme_minimal() +
  labs(
    x = element_blank(),
    y = "Frequency"
  )
```

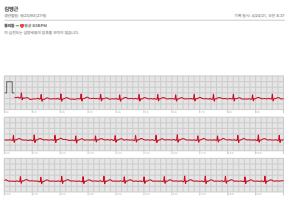
A.2 ECG Results



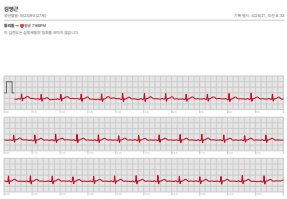
(a) (1,1),House blend: 89 BPM



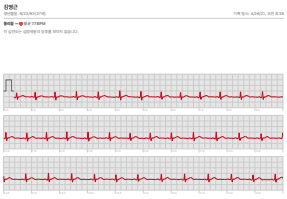
(b) (1,2),Water: 78 BPM



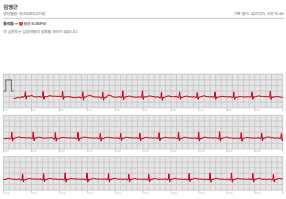
(c) (1,3),Sumatra: 83 BPM



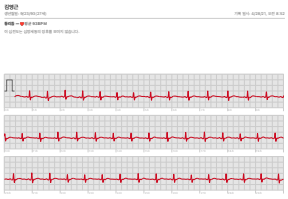
(d) (1,4),Decaf: 79 BPM



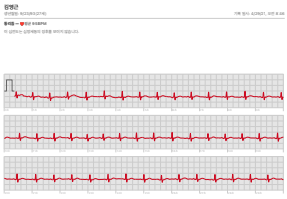
(e) (2,1),Water: 77 BPM



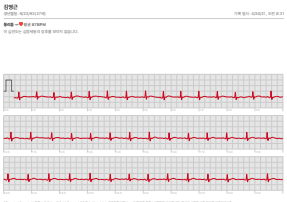
(f) (2,2),Sumatra: 83 BPM



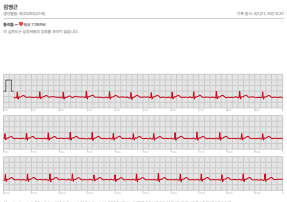
(g) (2,3),Decaf: 93 BPM



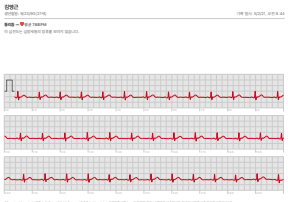
(h) (2,4),House blend: 95 BPM



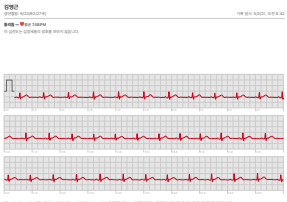
(i) (3,1),Sumatra: 87 BPM



(j) (3,2),Decaf: 77 BPM

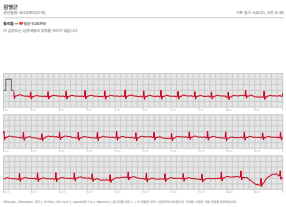


(k) (3,3),House blend: 78 BPM



(l) (3,4),Water: 74 BPM

(m) (4,1),Decaf: 77 BPM



(n) (4,2),House blend: 92 BPM

(o) (4,3),Water: 71 BPM

(p) (4,4),Sumatra: 76 BPM

Figure 3: Electrocardiogram after drinking Coffee